



**UNIVERSIDADE FEDERAL DE SERGIPE**  
**CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA**  
**DEPARTAMENTO DE ESTATÍSTICA E CIÊNCIAS ATUARIAIS**

**Gabrielle Monteiro Souza**

**ESTUDO DO ÍNDICE DE DESENVOLVIMENTO HUMANO DA  
POPULAÇÃO SERGIPANA EM RELAÇÃO À PRODUTIVIDADE DE  
LEITE POR MEIO DE TÉCNICAS MULTIVARIADAS**

**São Cristóvão – SE**

**2015**

**Gabrielle Monteiro Souza**

**ESTUDO DO ÍNDICE DE DESENVOLVIMENTO HUMANO DA  
POPULAÇÃO SERGIPANA EM RELAÇÃO À PRODUTIVIDADE DE  
LEITE POR MEIO DE TÉCNICAS MULTIVARIADAS**

**Monografia apresentada ao  
Departamento de Estatística e Ciências  
Atuariais da Universidade Federal de  
Sergipe, como parte dos requisitos para  
obtenção do grau de Bacharel em  
Estatística.**

**Orientador (a): Amanda da Silva Lira**

**São Cristóvão – SE**

**2015**

**Gabrielle Monteiro Souza**

**ESTUDO DO ÍNDICE DE DESENVOLVIMENTO HUMANO DA  
POPULAÇÃO SERGIPANA EM RELAÇÃO À PRODUTIVIDADE DE  
LEITE POR MEIO DE TÉCNICAS MULTIVARIADAS**

**Monografia apresentada ao  
Departamento de Estatística e Ciências  
Atuariais da Universidade Federal de  
Sergipe, como um dos pré-requisitos  
para obtenção do grau de Bacharel em  
Estatística.**

\_\_\_\_/\_\_\_\_/\_\_\_\_

**Banca Examinadora:**

---

**Prof.<sup>a</sup> Amanda da Silva Lira**  
**Orientadora**

---

**Prof.<sup>a</sup> Vanessa Kelly dos Santos**  
**1º Examinador**

---

**Prof.<sup>a</sup> Ana Flávia Menezes Santos**  
**2º Examinador**

## AGRADECIMENTOS

Não é uma tarefa fácil fazer esses agradecimentos, são tantos a agradecer, tantos que compartilharam dessa jornada e tantos que ainda caminham comigo nessa luta. Injusto seria citar alguns e, no momento, esquecer de outros não menos importantes. Por isso vai minha enorme gratidão aos colegas e amigos que estiveram em salas de aula, nos corredores, no DECAT e até mesmo aos que sempre apoiaram mesmo estando distantes.

Também passei pelos ensinamentos de muitos professores e a cada um de vocês dedico um espaço dos meus agradecimentos por toda dedicação e aprendizado passados. De uma forma mais carinhosa, agradeço a Amanda, de quem fui aluna, monitora e agora orientanda. Em especial a você, pela paciência e prontidão, mas principalmente pelo jeito doce com que trata as pessoas, o que foi determinante na minha escolha para orientador.

Meus mais sinceros agradecimentos vão às pessoas mais importantes da minha vida, minha família! A minha linda vovó Maria, aos meus avós Raquel e Tutu (em memória), tios e primos. Aos meus irmãos, Danielle e Rodolfo, cada um com seu jeito particular e especial de ser. Aos meus pais, não existe palavras de agradecimento, minha mãe, Conceição, minha amiga e conselheira, está sempre disposta a me confortar e sempre acreditou que este momento chegaria. Meu pai, Alvaci, é uma pessoa singular e de um caráter, comprometimento, calma, inteligência e tantos outros adjetivos sem igual. O que eu posso dizer ao meu querido pai, Roberto Carlos diz na canção “Esses seus cabelos brancos, bonitos...Sua voz macia me acalma e me diz muito mais do que eu digo, me calando fundo na alma, meu querido, meu velho, meu amigo”.

Agradeço a Deus, por tudo que me proporcionou e proporciona, pelo o que conquistei e o que tenho a conquistar. Sou agradecida por cada momento que me fez perceber sua presença em minha vida. Sou grata ainda por ter colocado em meu caminho uma pessoa mais que especial, a pessoa que me conquistou com seu jeito sério, mas brincalhão, determinado e sincero. Victor entrou na minha vida de forma inesperada e me mostrou que se pode conquistar sonhos com muita luta. A você, meu amigo, companheiro, namorado e marido, agradeço a compreensão, dedicação e amor que sempre me ofereceu. Por fim, agradeço a todos que, de alguma forma, contribuíram para a realização deste trabalho.

*“Eu queria voar e de tudo esquecer. No infinito morar e tranquilo viver.”*

Pe. Soares

## RESUMO

A bovinocultura e a pecuária leiteira são atividades que vem ganhando cada vez mais espaço na economia sergipana, ao mesmo tempo percebe-se que o IDH nos municípios também vem apresentando crescimento em seus índices. Diante disso, o presente trabalho objetiva avaliar possíveis influências que o crescimento na pecuária leiteira possa estar refletindo sobre os índices do IDHM das últimas três décadas. Com os dados disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e Programa das Nações Unidas para o Desenvolvimento (PNUD), foi possível organizar o banco de dados com as variáveis de efetivo bovino, produção leiteira, valor da produção, populações urbana e rural e o Índice de Desenvolvimento Humano Municipal (Longevidade, Renda e Educação) nos anos 1991, 2000 e 2010. Com os dados executaram-se técnicas como correlação e testes que avaliam a adequabilidade da aplicação da Análise Fatorial, uma técnica multivariada que permite identificar fatores para explicar a relação entre as variáveis em estudo. Outra técnica multivariada foi realizada a fim de verificar quais localidades possuem características similares entre si, a Análise de Agrupamentos. Ao final, os resultados mostraram que não há relação entre o crescimento do IDHM e a pecuária leiteira no estado de Sergipe, de forma que o crescimento de um não depende do crescimento do outro.

**Palavras-chaves:** Desenvolvimento humano; Pecuária leiteira; Análise fatorial; Análise de agrupamentos.

## **ABSTRACT**

The cattle and dairy farming are activities that are gaining more space in Sergipe's economy at the same time it's noticeable that the IDH (Human Development Index) in the municipalities has also been showing growth in their indexes. Therefore, this present study aims to evaluate possible influences that the growth in dairy farming can be reflecting on IDHM's indexes of the last three decades. With the data provided by the Brazilian Institute of Geography and Statistics (IBGE) and United Nations Development Program (PNUD), it was possible to organize the database with the effective bovine variables, milk production, value of production, urban and rural populations and the Municipal Human Development Index (longevity, income and education) in the years 1991, 2000 and 2010. With the data, it was executed techniques such as correlation and tests that evaluate the suitability of the application of Factor Analysis, a multivariate technique that allows to identify factors to explain the relationship between the variables under study. Another multivariate technique was performed in order to determine which localities have similar characteristics to each other, the Cluster Analysis. Finally, the results showed that there is no relationship between the growth of IDHM and dairy farming in the state of Sergipe, in a way that the growth of one does not affect the other.

**Key-words:** Human development; Dairy cattle; Factor analysis; Cluster analysis.

## LISTA DE QUADROS

Quadro 1.	Classificação do IDH.....	16
-----------	---------------------------	----

## LISTA DE FIGURAS

Figura 1.	Determinação dos fatores para os anos 1991, 2000 e 2010.....	30
Figura 2.	Dendrograma dos municípios de Sergipe em relação às variáveis dos IDHMs e da bovinocultura leiteira para o ano de 1991.....	34
Figura 3.	Dendrograma dos municípios de Sergipe em relação às variáveis dos IDHMs e da bovinocultura leiteira para o ano de 2000.....	35
Figura 4.	Dendrograma dos municípios de Sergipe em relação às variáveis dos IDHMs e da bovinocultura leiteira para o ano de 2010.....	36

## LISTA DE TABELAS

Tabela 1.	Correlação de Pearson para as variáveis do Teste 1 no ano de 1991.....	25
Tabela 2.	Correlação de Pearson para as variáveis do Teste 1 no ano de 2000.....	26
Tabela 3.	Correlação de Pearson para as variáveis do Teste 1 no ano de 2010.....	27
Tabela 4.	Autovalores e testes de KMO e Bartlett no Teste 1.....	28
Tabela 5.	Valores da diagonal principal da matriz de anti-imagem do Teste 1.....	28
Tabela 6.	Autovalores e testes de KMO e Bartlett do Teste 2.....	29
Tabela 7.	Valores da diagonal principal da matriz de anti-imagem do Teste 2.....	29
Tabela 8.	Matriz dos componentes rotacionados para o Teste 2.....	30
Tabela 9.	Coeficiente de correlação cofenética para os anos 1991, 2000 e 2010.....	31



## **LISTA DE ABREVIATURAS E SIGLAS**

PIB-	Produto Interno Bruto
IDH-	Índice de Desenvolvimento Humano
ONU-	Organização das Nações Unidas
PNUD-	Programa das Nações Unidas para o Desenvolvimento
IDHM-	Índice de Desenvolvimento Humano Municipal
KMO-	Kaiser-Meyer-Olkin
IBGE-	Instituto Brasileiro de Geografia e Estatística
SEBRAE-	Serviço Brasileiro de Apoio às Micro e Pequenas Empresas
PPM-	Pesquisa Pecuária Municipal
SIDRA-	Sistema IBGE de Recuperação Automática
IDHM-L-	Índice de Desenvolvimento Humano Municipal para Longevidade
IDHM-E-	Índice de Desenvolvimento Humano Municipal para Educação
IDHM-R-	Índice de Desenvolvimento Humano Municipal para Renda
MSA-	Medida de Adequação de Amostragem

## SUMÁRIO

1.	<b>INTRODUÇÃO.....</b>	11
2.	<b>OBJETIVOS.....</b>	13
2.1.	OBJETIVO GERAL.....	13
2.2.	OBJETIVOS ESPECÍFICOS.....	13
3.	<b>JUSTIFICATIVA.....</b>	14
4.	<b>REVISÃO LITERÁRIA.....</b>	15
4.1.	DESENVOLVIMENTO HUMANO E PECUÁRIA.....	15
4.1.1.	<b>Desenvolvimento Humano.....</b>	15
4.1.2.	<b>Pecuária.....</b>	16
5.	<b>METODOLOGIA.....</b>	18
5.1.	ANÁLISE MULTIVARIADA.....	19
5.1.1.	<b>Análise Fatorial.....</b>	19
5.1.1.1.	Correlação Linear.....	20
5.1.1.2.	Testes KMO e Bartlett.....	20
5.1.2.	<b>Análise de Agrupamentos.....</b>	21
5.1.2.1.	Medidas de Similaridade e Dissimilaridade.....	22
5.1.2.2.	Dendrograma.....	23
6.	<b>RESULTADOS E DISCUSSÕES.....</b>	24
7.	<b>CONCLUSÕES.....</b>	37
	<b>BIBLIOGRAFIA.....</b>	39
	<b>APÊNDICE A.....</b>	43
	<b>ANEXO A.....</b>	44

## 1 INTRODUÇÃO

O surgimento dos indicadores sociais, segundo Ottoni (2006), se deu a partir do descompasso existente entre os indicadores econômicos e as condições sociais da população ao longo do século XX. Nessa época o desenvolvimento social se media por meio do Produto Interno Bruto (PIB), mas essa medida não era a mais eficiente no determinante do bem-estar social e começou a ser discutida.

O crescente interesse pelas atividades em políticas públicas fez surgir a necessidade de reavaliar e criar instrumentos que viessem a apontar um desenvolvimento social mais condizente com a realidade. Com esse propósito, um dos indicadores criados foi o Índice de Desenvolvimento Humano (IDH) no início dos anos 1990 para suprir as deficiências de outros indicadores de caráter apenas econômico, como afirmam Guimarães e Jannuzzi (2005).

O IDH é um indicador utilizado pela Organização das Nações Unidas (ONU) para avaliar o nível do desenvolvimento humano de países no mundo, este índice é coletado pelo Programa das Nações Unidas para o Desenvolvimento (PNUD). Sua finalidade é a de medir, resumidamente, a qualidade de vida das pessoas, é o que define o PNUD.

Ainda seguindo as definições do PNUD, em meio aos vários índices existentes, o IDH é uma medida sintética do desenvolvimento humano de um país ou região. Sua metodologia abrange três dimensões básicas: educação, longevidade e renda. Cada uma destas dimensões fornece um índice, os quais, a partir de uma média aritmética simples, dão origem ao IDH Global.

O Índice de Desenvolvimento Humano Municipal – IDHM é uma extensão do IDH e segue o mesmo princípio do cálculo das três dimensões utilizadas para o cálculo do IDH Global. Ele é o mais apropriado para avaliar o desenvolvimento nos municípios, pois o IDHM faz adequação da metodologia utilizada no cálculo do índice Global à realidade brasileira. O índice municipal varia de 0 a 1, quanto mais próximo de 1 maior o desenvolvimento humano da localidade.

Segundo o Ranking de Desenvolvimento Humano, divulgado pelo PNUD, em 2010 Sergipe alcançou índice “médio” de desenvolvimento humano, com IDHM de 0,665. Nos anos 1991 e 2000, o estado apresentava índices “muito baixo” e “baixo”, respectivamente.

Com o passar dos anos, o IDH e IDHM têm apresentado aumentos substanciais no que diz respeito à qualidade de vida da sociedade. Os fatores econômicos e sociais que influenciam a melhoria na qualidade de vida são também os que determinam a ascensão destes índices. Para tanto, acredita-se que a pecuária bovina e leiteira, por se tratar de uma cultura antiga e que vem se desenvolvendo ao longo do tempo, pode estar ligada ao crescimento dos IDH's no estado.

O Brasil é considerado detentor de um dos maiores rebanhos bovinos comerciais do mundo. A bovinocultura de leite também intitula o país como um dos principais produtores mundiais, de acordo com a matéria de Fialho (2012) à Revista Agropecuária.

Apontada como uma das principais motivações da colonização de Sergipe, a pecuária ainda é uma atividade de grande importância no estado. A bovinocultura se insere como a principal atividade de subsistência e está introduzida em todos os 75 municípios de Sergipe.

O estado atualmente está composto por 8 (oito) territórios de planejamento: Alto Sertão Sergipano, Leste Sergipano, Médio Sertão Sergipano, Agreste Central Sergipano, Baixo São Francisco Sergipano, Grande Aracaju, Centro-Sul Sergipano e Sul Sergipano (Anexo A, pág. 44). Desses territórios, a bovinocultura leiteira está inserida, principalmente, no alto sertão sergipano, centro sul, agreste central e médio sertão.

O leite é o principal produto extraído do animal e se evidencia no alto sertão sergipano. Essa região concentra as maiores bacias leiteiras do estado, com destaque para Nossa Senhora da Glória. A região ainda é conhecida pela produção dos derivados do leite, a exemplo do queijo, da manteiga e do requeijão.

Contudo, o objetivo deste trabalho é verificar as possíveis relações que possam existir entre o IDHM, os efetivos do rebanho e a produção de leite dos municípios do estado de Sergipe que possam justificar seus crescimentos contínuos.

## **2 OBJETIVOS**

### **2.1 Geral**

O trabalho em foco tem por objetivo geral conferir se existe relação entre o Índice de Desenvolvimento Humano Municipal (IDHM): Longevidade, Educação e Renda; os efetivos de rebanho e produtividade leiteira e a população rural e urbana dos municípios do estado de Sergipe nos três últimos anos censitários, 1991, 2000 e 2010. Com isso, o presente estudo tem a finalidade de apresentar à população sergipana a importância das atividades relacionadas à bovinocultura nos índices do IDHM que avaliam o desenvolvimento humano dos municípios.

### **2.2 Específicos**

Com o propósito de alcançar o objetivo geral deste trabalho, foram determinados os seguintes objetivos específicos:

- Obter os dados referentes ao IDHM, efetivos de bovino e produção leiteira e população de cada município sergipano nos três anos que a pesquisa abrange;
- Realizar testes estatísticos que validem a possibilidade de execução dos estudos por meio da Análise Fatorial, tais como: a correlação linear, testes de Kaiser-Meyer-Olkin (KMO) e esfericidade de Bartlett.
- Aplicar análise de agrupamentos com a finalidade de formar grupos, os quais serão compostos pelos municípios com características equivalentes entre si.
- Dispor as análises em forma de tabelas e figuras, de modo que possa facilitar a visibilidade dos resultados.

### **3 JUSTIFICATIVA**

Partindo da ideia da professora Msc. Amanda da Silva Lira, em sua dissertação apresentada à Universidade Federal Rural de Pernambuco em 2010, que fez o estudo do desenvolvimento humano e da pecuária leiteira para o estado de Pernambuco, surgiu o interesse e a curiosidade em aplicar o mesmo estudo para os municípios de Sergipe tendo em vista a importância da atividade pecuarista e o aumento nos níveis de desenvolvimento humano no estado.

A pecuária é uma atividade relevante para a população sergipana e vem evoluindo com o tempo tanto em seu efetivo quanto em sua produção leiteira, que é o principal produto extraído da pecuária bovina que o estado produz. Sergipe possui 75 municípios e em todos eles a atividade bovina está presente.

Ao mesmo tempo em que a bovinocultura cresce, o IDH do estado e dos municípios vem apresentando melhorias ao longo dos anos. Devido a esse crescimento, a pesquisa ora apresentada se justifica na necessidade de investigar se a criação bovina tem, de alguma forma, contribuído para a ascensão do desenvolvimento nos municípios de Sergipe.

Para explorar esse tema serão utilizados dois métodos estatísticos multivariados, a análise fatorial e de agrupamentos. O método da análise fatorial tem por característica separar componentes por meio de fatores que apresentem correlação entre si. A análise de agrupamentos objetiva formar grupos compostos por elementos com características similares. Dado isso, demonstraremos aqui testes e métodos que mostrem se há relação entre as atividades da bovinocultura e o aumento do desenvolvimento social.

## **4 REVISÃO LITERARIA**

### **4.1 Desenvolvimento Humano e Pecuária**

Esta seção tratará dos aspectos teóricos referentes ao trabalho, sendo dividido em duas subseções, que abordam respectivamente: Desenvolvimento Humano e Pecuária.

#### **4.1.1 Desenvolvimento Humano**

Durante um longo período o desenvolvimento de uma nação era baseado apenas em medidas econômicas, a exemplo do Produto Interno Bruto – PIB que calcula a soma das riquezas geradas pelo país. Com esse pensamento, quanto maior fosse o valor da produção de uma nação, mais desenvolvida ela era. Porém, o Programa das Nações Unidas para o Desenvolvimento (PNUD) sugeriu um novo indicador que visasse o desenvolvimento não apenas pelas riquezas, como também por dimensões essenciais à vida.

O conceito de desenvolvimento humano, de acordo com o PNUD, surgiu como um método de ampliação do bem-estar e oportunidades para as pessoas terem livre arbítrio para escolherem a vida que desejarem ter. Conceitua ainda que o desenvolvimento humano deve estar aliado ao aumento de capacidades e oportunidades das pessoas como um dos meios pelo qual elas possam viver a vida que desejarem, e não apenas no aumento de riquezas e renda como o único meio de bem-estar.

De acordo com Sen (2000, apud ROQUE 2009), o desenvolvimento consiste em eliminar as privações de liberdade que limitam as escolhas e as oportunidades das pessoas de exercer sua condição de agente. Para ele, o desenvolvimento só é formado quando não há privações de liberdades.

O Índice de Desenvolvimento Humano (IDH) foi publicado a primeira vez em 1990 com o propósito de ser uma medida genérica e sucinta do desenvolvimento social. Como define o PNUD, o IDH surgiu como uma forma de medir a qualidade de vida das pessoas dos países a partir de dimensões como a longevidade, a educação e a renda.

O IDH para os municípios – IDHM foi inserido no Brasil em 1998, ele é uma adaptação do IDH global à realidade local. Segundo Kieling (2014), o cálculo para o

IDHM é feito a partir de dados estatísticos fornecidos pelo IBGE, diferentemente do IDH global que é calculado através dos dados da agência da Organização das Nações Unidas (ONU) e outras agências que parametrizam os dados nacionais a fim de que haja comparabilidade entre os países.

Em sua formulação essas dimensões são calculadas de modo que seus valores estejam entre 0 e 1, indicando que quanto mais próximo de 1 estiver o índice, maior o desenvolvimento social da respectiva localidade. Em seu ranking, a PNUD classifica os municípios conforme o quadro abaixo:

Quadro 1: Classificação do IDH

<b>Classificação</b>	<b>Intervalo</b>
Muito Alto	0,80 - 1,00
Alto	0,70 - 0,79
Médio	0,60 - 0,69
Baixo	0,50 - 0,59
Muito Baixo	0,00 - 0,49

Fonte: Atlasbrasil

Com os dados utilizados nesta pesquisa temos que, para o ano 2010, o IDHM do município de Aracaju é o maior do estado com 0,770, a partir do Quadro 1 apresentado, verificamos que este município tem desenvolvimento humano considerado alto. Em contrapartida, a localidade com o menor índice é Poço Redondo com IDHM 0,529, o que significa dizer que a localidade possui um baixo desenvolvimento social.

#### 4.1.2 Pecuária

Sergipe é a menor unidade da federação situado na região Nordeste, conta com uma área de 21.918,493 km<sup>2</sup> distribuída entre 75 municípios e população de 2.068.017 habitantes em 2010 IBGE (2015). Atualmente o estado possui 8 Territórios de Planejamento que foram elaborados pelo governo estadual a partir de pesquisas feitas com a população e instituído pelo Decreto nº 24338 de 2007, como diz Vargas (2011). O processo de divisão foi realizado com base em vários aspectos, sobretudo os aspectos



culturais com objetivos e metas que justificassem a inserção da população em um determinado território.

Um dos importantes fatores econômicos que motivaram a colonização em Sergipe foi a pecuária. Essa cultura desenvolveu-se, principalmente, pelas condições territoriais propícias e pela necessidade de pouca mão de obra. Conforme afirmam Carvalho e Souto (2011) nota-se que a organização territorial, dentre outros aspectos sociais, sofreu grande influência da pecuária, da cana-de-açúcar e da cultura do algodão durante o período da colonização, no século XVI até o século XX.

Sergipe é um estado que tem a particularidade de possuir a pecuária bovina presente em todos seus municípios, destacam-se no ramo as regiões Centro Sul, Agreste Central, Médio Sertão e, principalmente, o Alto Sertão Sergipano como principal território nessa atividade.

Menezes (2012) relata que no Sertão Sergipano, apesar de a pecuária ser a principal atividade econômica, não era considerada uma cultura muito forte até meados do século XX. Devido aos constantes períodos de estiagem existentes nessa região, observaram-se que os prejuízos são mais intensos na agricultura, desse modo foram criadas políticas públicas que incentivassem a transformação de lavouras em pastagens intensificando uma outra atividade ligada à pecuária.

Na década de 1970, ocorreu, nos estabelecimentos pertencentes a camponeses do Sertão Sergipano, a expansão da pecuária leiteira. O pequeno estabelecimento rural (com até 50 ha), anteriormente ocupado com lavouras de milho, feijão, mandioca e algodão (consorciados), foi modificado, com a inserção de pastagens direcionadas ao gado leiteiro (MENEZES, 2012).

O município de Nossa Senhora da Glória abriga a maior bacia leiteira do estado e é conhecida também pelos produtos produzidos a partir do leite, como o queijo, manteiga e requeijão. Um levantamento feito pelo Serviço Brasileiro de Apoio às Micro e Pequenas Empresas – SEBRAE em 2014, a partir de dados do IBGE, apontou para um crescimento percentual de destaque no Nordeste na produção de leite do estado.

Os dados da Pesquisa Pecuária Municipal – PPM, feita pelo IBGE, mostram que o estado teve crescimento de 18,0% na produção de leite entre 1991 e 2000. Já de 2000 até 2010, o crescimento foi muito além, atingiu 157,6%, quando a produção que era de 115,1 milhões de litros em 2000 passou para 296,6 milhões de litros em 2010.

## 5 METODOLOGIA

Para a elaboração desse trabalho foi utilizado material de natureza descritiva, sendo utilizadas pesquisas bibliográficas e consultas a internet. O presente estudo busca avaliar a possível relação entre o IDHM e a pecuária bovina nos municípios que compõem o estado de Sergipe, para tanto foram obtidas informações nos anos censitários 1991, 2000 e 2010 referentes à pecuária: número de vacas ordenhadas, produção de leite em litros ( $10^3$ ), número de cabeças de gado e valor da produção em reais, inclusive o ano de 1991 e variáveis relacionadas à população: população urbana e rural do Sistema IBGE de Recuperação Automática (SIDRA). Do PNUD, foram coletadas informações sobre o IDH dos Municípios: IDHM, IDHM-L (Longevidade), IDHM-R (Renda) e IDHM-E (Educação) também para os anos 1991, 2000 e 2010.

Os dados foram analisados separadamente por ano, de modo que foram testadas inicialmente as variáveis do ano 1991, depois as variáveis de 2000 e por último as variáveis de 2010. A partir dos dados foi feita uma Análise Fatorial e, após a validação da possibilidade de aplicar a Análise Fatorial, foi utilizado o método de Cluster (Análise de Agrupamentos).

Primeiramente foi executado o cálculo da correlação linear de Pearson com o propósito de confirmar se há relação entre as variáveis dos IDHs (educação, renda e longevidade), as variáveis de população (urbana e rural) e as variáveis referentes à pecuária (de rebanho e produtividade leiteira). Para testar a correlação entre os pares das variáveis, seguindo orientações de Bezerra (2007), testou-se a significância de acordo com o teste “t”, com probabilidade de 5% ( $\alpha=0,05$ ).

O passo seguinte tratou da utilização de dois testes indicatórios da viabilidade ou não de proceder com a Análise Fatorial. Os testes foram o de Kaiser-Meyer-Olkin (KMO), que mensura o grau de explicação dos dados e o teste de Esfericidade de Bartlett que indica se há relação satisfatória entre as variáveis, como orienta Bezerra (2007).

O método de extração de fatores aplicado foi o de componentes principais, que permite extrair os fatores que explicam o máximo da variância dos dados (BEZERRA, 2007). Para aumentar a capacidade explicativa dos fatores obtidos com o método de componentes principais, optou-se por fazer uma readaptação dos autovalores a partir da rotação Varimax.

O teste de validação da aplicação da Análise Fatorial foi realizado duas vezes, e a repetição foi chamada de Teste 1 e Teste 2. O segundo teste foi executado na tentativa de aumentar o poder de explicação dos fatores devido o teste de KMO ter apresentado valores, apesar de suficientes para a aplicação da Análise Fatorial, de baixo poder explicativo entre as variáveis e os fatores dos dados originais.

Por último, foi executada a separação dos dados em grupos por meio da análise de agrupamentos, com a finalidade de observar quais os municípios possuem maiores afinidades entre si, no tocante à relação entre o IDHM e a pecuária bovina e leiteira. Para tanto, escolheu-se como medida de avaliar a semelhança, a distância euclidiana e a técnica hierárquica foi a de ligação completa. Procedeu-se ainda com a análise a partir de dendrogramas para uma melhor visualização dos grupos formados mediante as técnicas descritas.

As análises foram realizadas a partir do software Statistical Package for Social Sciences (SPSS) versão 17.0 e R-project versão 3.2.1.

## **5.1. Análise Multivariada**

É o conjunto de técnicas estatísticas que tem por finalidade de analisar múltiplas variáveis simultaneamente, em que o resultado de cada variável leva em conta o efeito das demais variáveis. “É somente por meio de técnicas multivariadas que essas múltiplas relações podem ser adequadamente examinadas para se obter uma compreensão mais completa e realista na tomada de decisões” (HAIR Jr. et al., 2005).

A análise multivariada compreende diversas técnicas para análise de dados, tais como a análise fatorial, regressão múltipla, agrupamentos, análise de correspondência, dentre outras. Aqui, serão tratadas apenas das análises fatorial e de agrupamentos.

### **5.1.1 Análise Fatorial**

É uma técnica estatística onde o principal objetivo é explicar a correlação entre variáveis de um dado conjunto. A análise fatorial tenta representar um conjunto de dados por meio de um número menor de dimensões, que são os chamados “fatores”.

A literatura de Bezerra (2007) afirma que cada variável pode ser independente entre si, neste caso a dimensão será de acordo com o número de variáveis existentes. No

entanto, se houver relações de dependência entre as variáveis, existirá agrupamentos que, consequentemente, gerarão um número de fatores menor que o número de variáveis.

Contudo, a análise fatorial é utilizada para discriminar fatores que possuam um alto grau de explicação da variabilidade dos dados originais.

### 5.1.2 Correlação Linear

A correlação linear ou correlação de Pearson –  $r$ , é um coeficiente utilizado para estudar o grau de associação entre duas variáveis, uma dependente e a outra independente. Este coeficiente, além do grau de intensidade da relação entre duas variáveis, mede também em qual direção está essa relação.

Segundo Barbetta (2002), o valor  $r$  do coeficiente de Pearson é um valor dentro do intervalo de -1 a 1. Quanto mais próximo  $r$  estiver de 1 (ou -1), mais forte é a correlação entre as variáveis. Se o valor da correlação estiver próximo de zero, significa que não há correlação dos dados. A correlação, quanto ao seu sentido, pode ser negativa ou positiva e quanto a sua força, pode ser fraca, moderada ou forte.

O autor descreve ainda uma fórmula, por meio da qual é feito o cálculo do coeficiente  $r$ :

$$r = \frac{n \cdot \sum(X \cdot Y) - (\sum X) \cdot (\sum Y)}{\sqrt{n \cdot \sum X^2 - (\sum X)^2} \cdot \sqrt{n \cdot \sum Y^2 - (\sum Y)^2}}$$

em que,  $r$  representa o coeficiente de correlação de Pearson,  $n$  é o tamanho amostral e  $X$  e  $Y$  são as variáveis.

### 5.1.3 Testes KMO e Bartlett

Ambos os critérios são utilizados para a validação dos dados para a aplicação da Análise Fatorial. O teste proposto por Kaiser-Meyer-Olkin é uma medida representada por um índice Measure of Sampling Adequacy (MSA) que é de bastante importância e testa a consistência geral dos dados.

Cruz e Topa (2009) afirma que, a metodologia é verificada a partir da proximidade da matriz de correlação inversa com a matriz diagonal, e se fundamenta em

comparar os valores dos coeficientes de correlação linear observados com os valores dos coeficientes de correlação parcial.

Bezerra (2007) diz que os resultados vão variar de 0 a 1, de forma que os valores entre 0,5 até 1 indicam adequação à aplicação da análise fatorial. Já os valores abaixo de 0,5 não são adequados à aplicação da análise fatorial.

O cálculo é executado pela fórmula:

$$MSA: \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2}$$

de modo que:

$r_{ij}$  – é o coeficiente de correlação simples entre as variáveis  $x_i$  e  $x_j$ .

$a_{ij}$  – é o coeficiente de correlação parcial  $x_i$  e  $x_j$ , ou seja, os quadrados dos valores que ficam fora da diagonal da matriz de anti-imagem.

O teste de esfericidade de Bartlett testa a hipótese de que a matriz de correlação é uma matriz identidade, ou seja, testa a hipótese de não haver correlação entre os dados. Caso a hipótese nula testada seja rejeitada, a um nível de significância de 5%, pode-se aplicar a análise fatorial, segundo Ferreira Junior (2004, apud CRUZ e TOPA 2009).

Marques (2010) mostra a estatística de prova dada por:

$$X^2 = - \left[ (n - 1) - \frac{2p + 5}{6} \right] \ln|R|$$

em que possui distribuição qui-quadrado com  $v = \frac{p(p-1)}{2}$  graus de liberdade, com:

$n$  = tamanho da amostra.

$p$  = número de variáveis.

$|R|$  = determinante da matriz de correlação.

#### 5.1.4 Análise de Agrupamentos

Essa é uma técnica analítica que tem por finalidade formar grupos de acordo com as características dos objetos. Também conhecida como análise de cluster, a análise de agrupamentos se baseia na similaridade entre as características dos objetos para a formação dos grupos, de modo que o resultado final é de homogeneidade dentro dos grupos e heterogeneidade entre os grupos.

Hair Jr. et al. (2005) define que, no mínimo, são três os passos que precisam ser avaliados numa análise de agrupamentos. O primeiro é a medida de alguma forma de similaridade, o segundo é como o agrupamento é formado e o terceiro é a determinação dos grupos.

#### 5.1.4.1 Medidas de Similaridade e Dissimilaridade

Na aplicação da análise de agrupamentos, deve-se fazer a escolha de qual coeficiente de definição dos grupos mais parecidos utilizar, coeficiente de similaridade ou de dissimilaridade. Com o coeficiente de similaridade, quanto maior o valor observado, mais similares são os objetos. Já na medida de dissimilaridade, quanto maior o valor observado, menos similares serão os objetos (VICINI, 2005).

Dentre as medidas de similaridade, as mais comuns são a correlação de Pearson e a similaridade do cosseno e dentre as de dissimilaridade, as mais conhecidas são as distâncias de Manhattan e, a mais comumente utilizada, euclidiana.

“A distância euclidiana é interpretada como a distância entre dois indivíduos, cujas posições são determinadas em relação às suas coordenadas, definidas com referência a um grupo de eixos cartesianos, dos quais possuem ângulos retos entre si”. Clifford & Stephenson (1975, apud MEYER, 2002). Considerando um par de elementos amostrais  $i$  e  $j$ , sua fórmula é descrita da seguinte forma:

$$d_{ij} = [(x_{ia} - x_{ja})^2 + (x_{ib} - x_{jb})^2]^{\frac{1}{2}},$$

em que  $a$  e  $b$  são as características,  $x_{ia}$  e  $x_{ib}$  representam as características do indivíduo  $i$  e  $x_{ja}$  e  $x_{jb}$  representam as características do indivíduo  $j$ .

Na modalidade dissimilaridade ainda temos dois processos de aglomeração, o hierárquico e o não-hierárquico. O método hierárquico é composto por dois tipos de agrupamentos, o divisivo e o aglomerativo. Este último ainda está subdividido em: métodos de encadeamento (ligação simples, ligação completa, ligação média), métodos de variância (Ward) e métodos de centroide.

Neste trabalho optou-se por utilizar o método de encadeamento por ligação completa, por este motivo, dos métodos já citados, este será o método de ligação detalhado.

Também conhecido como método do vizinho mais distante, a ligação completa aloca em cada grupo os indivíduos que menos se parecem, daí todos os pares são formados com um membro de cada grupo, dessa forma a similaridade entre dois grupos é determinada pelos objetos de cada grupo menos semelhantes, segundo Bussab et al. (1990).

[...] **métodos não-hierárquicos**, que se caracterizam por procurar maximizar a homogeneidade intragrupo, sem considerar a hierarquia entre grupos. Estes métodos costumam ser chamados de k médias ou k-means clustering. Kmeans clustering compreendem o limiar seqüencial, o limiar paralelo e o particionamento otimizador. (MALHOTRA, 2001 apud, VICINI, 2005, p.530 e 531).

Os métodos aglomerativos seguem basicamente uma mesma rotina, faz-se uma matriz das distâncias entre as variáveis e depois se verifica as uniões entre os grupos através da construção de um dendrograma, que possui uma escala para que sejam observados os níveis. A validação das informações, obtidas pelos processos de aglomeração, pode ser feita a partir do cálculo da correlação cofenética. Esse é o processo pelo qual se verifica a correlação entre a matriz de similaridade e a matriz de correlação cofenética e dá origem ao coeficiente de correlação cofenética. Quanto mais se aproximar de 1, é considerada forte e apropriada e, de acordo com Bussab et al. (1990), um coeficiente de correlação cofenética é considerado um bom ajuste a partir de 0,8.

#### 5.1.4.2 Dendrograma

É uma representação gráfica, em forma de árvore, característica da análise de agrupamento hierárquica, ela é considerada uma simplificação da matriz de similaridades. Bussab et al. (1990) afirma que “a escala vertical à esquerda, indica o nível de similaridade. No eixo horizontal são marcados os objetos, numa ordem conveniente, as linhas verticais partindo dos objetos têm altura correspondente ao nível em que os objetos são considerados semelhantes”.

## 6 RESULTADOS E DISCUSSÕES

Uma primeira análise foi realizada com o propósito de verificar a adequabilidade da utilização da Análise Fatorial a partir das variáveis dispostas, que foi chamada de Teste 1. A começar pela correlação entre as variáveis IDHM, IDHM-R, IDHM-L e IDHM-E e as variáveis relativas à produção de leite, valor da produção, vacas ordenhadas e os efetivos de rebanho bovino nos municípios do estado de Sergipe, nos anos 1991, 2000 e 2010. Essa correlação não foi satisfatória, apresentando, em sua maioria, valores menores que 0,40 e relações negativas entre os IDHMs e as variáveis referentes à pecuária (Tabelas 1, 2 e 3).

O segundo passo foi avaliar, por meio da Measure of Sampling Adequacy – MSA (Medida de Adequação de Amostragem) de KMO, o grau de explicação dos dados que não pode ser inferior a 0,50. Nesse caso, foram obtidos valores acima de 0,50 nos três anos, sendo que em 2000 e 2010 o MSA foi pouco acima do limite, caracterizando um baixo poder de explicação entre as variáveis e os fatores extraídos. Ainda podemos visualizar o teste de esfericidade de Bartlett, cuja significância apresentada foi de 0,000 nos três anos, consequentemente podemos dizer que é possível a aplicação da Análise Fatorial devido o valor de Significância ser inferior a 0,05. No entanto, mesmo o teste de Bartlett afirmando a possibilidade da utilização da Análise Fatorial, optou-se por tentar aumentar o poder de explicação dos fatores, com a exclusão de algumas variáveis (Tabela 4). Nas mesmas tabelas pode-se verificar ainda a variação total explicada, que indica o percentual de explicação dos fatores selecionados. Em 1991, dois fatores foram selecionados e eles, juntos, explicam 73,13% da variância. Em 2000, 77,07% da variância é explicada também por dois fatores. E em 2010 três fatores foram capazes de explicar 85,07% da variância dos dados.

O passo seguinte foi verificar as variáveis a serem retiradas da análise, com a ajuda de uma matriz anti-imagem. Nela, verifica-se a diagonal principal da tabela de Correlação anti-imagem, a variável que estiver abaixo de 0,50 é passível de exclusão da análise por ser pouco explicativa. A Tabela 5 mostra algumas variáveis suscetíveis a corte, no entanto foi preferível retirar apenas a que tem menor valor e verificar se, sem sua influência, as outras variáveis que receberam indicativo de corte aumentam suas capacidades de explicação, caso não haja evolução nesse sentido, outra(s) variável(is) será(ão) retirada(s) até que haja um bom conjunto de variáveis explicativas.



Tabela 1. Correlação de Pearson para as variáveis do Teste 1 no ano de 1991

Variáveis	IDHM	IDHM-Renda	IDHM-Longevidade	IDHM-Educação	Nº Bovinos-Cabeças	Produção de Leite-Litro	Valor da Produção-Reais	Nº Vacas Ordenhadas	Pop Urbana
IDHM-Renda	0,767**								
IDHM-Longevidade	0,545*	0,446							
IDHM-Educação	0,970**	0,663*	0,386						
Nº Bovinos-Cabeças	-0,302	0,011	-0,216	-0,350					
Produção de Leite-Litro	-0,370	-0,141	-0,205	-0,388	0,590*				
Valor da Produção-Reais	-0,378	-0,139	-0,215	-0,396	0,635*	0,992**			
Nº Vacas Ordenhadas	-0,406	-0,048	-0,253	-0,449	0,774**	0,867**	0,888**		
Pop Urbana	0,504*	0,643*	0,268	0,492*	-0,036	-0,072	-0,071	-0,070	
Pop Rural	-0,245	0,039	-0,194	-0,290	0,803**	0,398	0,455*	0,615*	-0,039

\* Significativo a 5% de probabilidade

\*\* Significativo a 1% de probabilidade

Tabela 2. Correlação de Pearson para as variáveis do Teste 1 no ano de 2000

Variáveis	IDHM	IDHM-Renda	IDHM-Longevidade	IDHM-Educação	Nº Bovinos-Cabeças	Produção de Leite-Litro	Valor da Produção-R\$	Nº Vacas Ordenhadas	Pop Urbana
IDHM-Renda	0,831**								
IDHM-Longevidade	0,577*	0,501*							
IDHM-Educação	0,958**	0,682*	0,393						
Nº Bovinos-Cabeças	-0,366	-0,046	-0,155	-0,459*					
Produção de Leite-Litro	-0,316	-0,094	-0,188	-0,362	0,752**				
Valor da Produção-R\$	-0,288	-0,041	-0,127	-0,357	0,804**	0,981**			
Nº Vacas Ordenhadas	-0,345	-0,081	-0,204	-0,405	0,882**	0,957**	0,966**		
Pop Urbana	0,564*	0,671*	0,334	0,494*	-0,056	-0,063	-0,054	-0,065	
Pop Rural	-0,347	-0,082	-0,085	-0,437	0,779**	0,459*	0,501*	0,625*	-0,051

\* Significativo a 5% de probabilidade

\*\* Significativo a 1% de probabilidade

Tabela 3. Correlação de Pearson para as variáveis do Teste 1 no ano de 2010

Variáveis	IDHM	IDHM-Renda	IDHM-Longevidade	IDHM-Educação	Nº Bovinos-Cabeças	Produção de Leite-Litro	Valor da Produção-R\$	Nº Vacas Ordenhadas	Pop Urbana
IDHM-Renda	0,865**								
IDHM-Longevidade	0,501*	0,459*							
IDHM-Educação	0,917**	0,643*	0,218						
Nº Bovinos-Cabeças	-0,292	-0,056	0,094	-0,451*					
Produção de Leite-Litro	-0,305	-0,194	-0,075	-0,340	0,590*				
Valor da Produção-R\$	-0,306	-0,193	-0,068	-0,345	0,608*	0,999**			
Nº Vacas Ordenhadas	-0,340	-0,159	-0,016	-0,434	0,818**	0,927**	0,934**		
Pop Urbana	0,637*	0,685*	0,388	0,501*	-0,050	-0,060	-0,059	-0,066	
Pop Rural	-0,192	-0,055	0,207	-0,331	0,784**	0,331	0,351	0,555*	-0,017

\* Significativo a 5% de probabilidade

\*\* Significativo a 1% de probabilidade

Tabela 4. Autovalores e testes de KMO e Bartlett no Teste 1

Ano	KMO	Bartlett (sig)	Autovalor	Nº de Fatores
1991	0,615	0,000	73,13%	2
2000	0,583	0,000	77,07%	2
2010	0,546	0,000	85,07%	3

Tabela 5. Valores da diagonal principal da matriz de anti-imagem do Teste 1

Variáveis	MSA		
	1991	2000	2010
IDHM	0,513	<b>0,476*</b>	<b>0,457*</b>
IDHM-Renda	<b>0,476*</b>	<b>0,384*</b>	<b>0,380*</b>
IDHM-Longevidade	<b>0,337*</b>	<b>0,251*</b>	<b>0,187*</b>
IDHM-Educação	0,508	<b>0,462*</b>	<b>0,420*</b>
Nº Bovinos-Cabeças	0,823	0,802	0,757
Produção de Leite-Litro	0,712	0,749	0,697
Valor da Produção-R\$	0,718	0,800	0,705
Nº Vacas Ordenhadas	0,925	0,809	0,859
Pop Urbana	<b>0,486*</b>	<b>0,487*</b>	0,588
Pop Rural	0,757	0,864	0,816

\* Valores menores que 0,5 podem ser retirados da análise

Ainda na Tabela 5, vê-se que em 1991 a indicação é para a retirada do IDHM-R, IDHM-L e da População urbana. Em 2000, mostravam-se pouco explicativas as variáveis IDHM, IDHM-R, IDHM-L, IDHM-E e a População urbana. Já em 2010, o indicativo era para excluir as variáveis IDHM, IDHM-R, IDHM-L e IDHM-E. Porém, a exclusão, à risca, das variáveis indicadas retiraria da análise as variáveis que são base para a pesquisa, que é a associação da pecuária com o índice de desenvolvimento humano dos municípios de Sergipe. Para a segunda análise, decidiu-se agir com cautela e retirar apenas a variável IDHM-L que, de acordo com a matriz de anti-imagem, é a de menor poder de explicação nos três anos. Dessa forma, se for obtido valores satisfatórios, será retirada apenas uma variável do estudo, caso contrário será repetido um outro teste retirando mais uma variável.

O Teste 2 seguiu os mesmos critérios do Teste 1, não necessitando a verificação dos coeficientes de correlação, pois essas se mantêm iguais às geradas no procedimento anterior. Logo, iniciou-se a segunda avaliação a partir do critério de KMO, que

anteriormente tinha revelado valores próximos ao mínimo aceitável, no novo teste apresentou aumento expressivo para os três anos, como pode ser visto na Tabela 6. Ainda na mesma tabela, vê-se que o teste de esfericidade de Bartlett continua confirmando a utilização da Análise Fatorial, indicando significância de 0,000 para os três anos e que o modelo continua sendo explicado pelos mesmos fatores, contudo esses fatores possuem maior poder de explicação. A variação total em 1991 permanece sendo explicada por dois fatores, que agora assume 77,90% do modelo. Em 2000, a variância é explicada por dois fatores que somam 81,89%. Em 2010 três fatores explicam 89,75% da variância. Com a retirada de apenas uma variável, a nova matriz de anti-imagem não indica a ser retirada mais nenhuma variável (Tabela 7).

Tabela 6. Autovalores e testes de KMO e Bartlett do Teste 2

Ano	KMO	Bartlett (sig)	Autovalor	Nº de Fatores
1991	0,702	0,000	77,90%	2
2000	0,712	0,000	81,89%	2
2010	0,698	0,000	89,75%	3

Tabela 7. Valores da diagonal principal da matriz de anti-imagem do Teste 2

Variáveis	MSA		
	1991	2000	2010
IDHM	0,608	0,588	0,597
IDHM-Renda	0,545	0,562	0,561
IDHM-Educação	0,644	0,610	0,627
Nº Bovinos-Cabeças	0,825	0,796	0,752
Produção de Leite-Litro	0,708	0,739	0,694
Valor da Produção-R\$	0,712	0,783	0,701
Nº Vacas Ordenhadas	0,899	0,793	0,857
Pop Urbana	0,625	0,799	0,928
Pop Rural	0,748	0,820	0,749

Ficou evidente que o segundo teste, com a retirada da variável de IDHM de longevidade, apresenta explicação melhor dos dados, consequentemente, confirma ser relevante a aplicação da análise fatorial. Na Figura 1 ficam visíveis os autovalores dos fatores que foram selecionados em cada ano, e as variáveis que compõem cada fator são mostrados na Tabela 8. Nessas representações visualiza-se em 1991 e em 2000 dois fatores selecionados, nos quais o Fator 1 é composto pelas variáveis relacionadas à pecuária bovina e a população rural e o fator 2 pelas variáveis de IDHM e a população

urbana. Em 2010, o primeiro fator relacionou os IDHM e a população urbana, o segundo fator ficou ligado à produção de leite, à quantidade de vacas ordenhadas e ao valor da produção, enquanto no terceiro fator ficaram o efetivo de bovinos e a população rural.

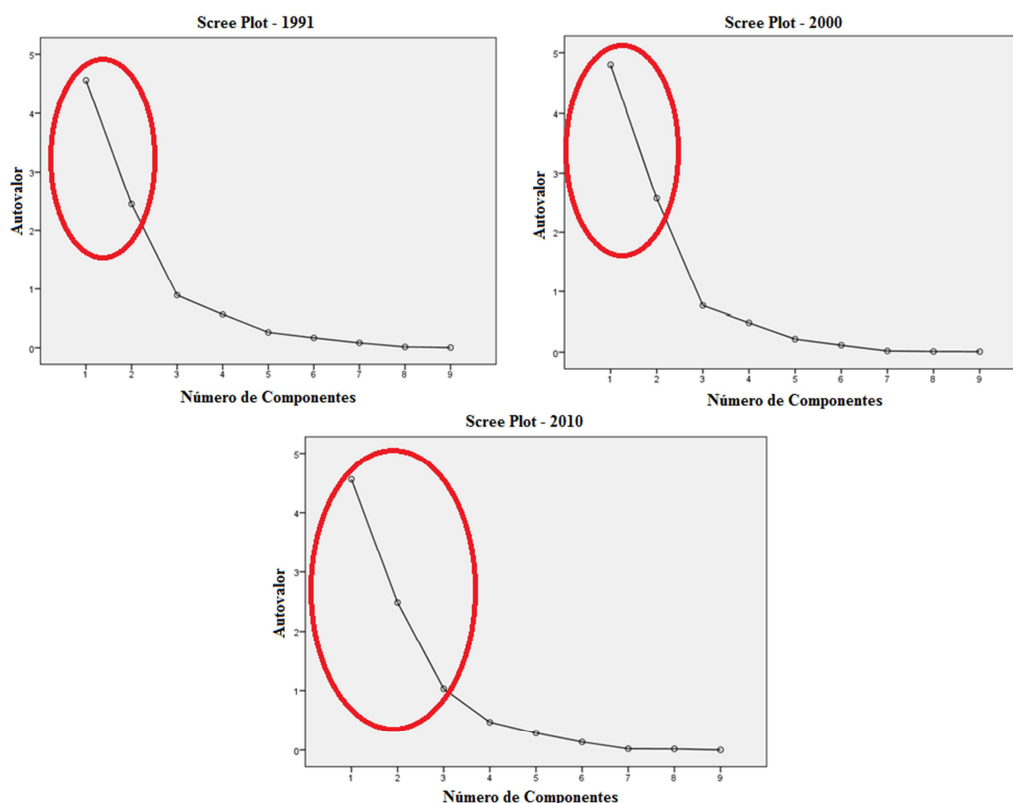


Figura 1 – Determinação dos fatores para os anos 1991, 2000 e 2010

Tabela 8. Matriz dos componentes rotacionados para o Teste 2

Variáveis	1991		2000		2010		
	Fator 1	Fator 2	Fator 1	Fator 2	Fator 1	Fator 2	Fator 3
IDHM	-0,308	0,897	-0,280	0,923	0,941	-0,184	-0,154
IDHM-Renda	0,056	0,914	0,035	0,926	0,917	-0,098	0,071
IDHM-Educação	-0,358	0,851	-0,374	0,843	0,809	-0,208	-0,354
Nº Bovinos-Cabeças	0,866	-0,040	0,922	-0,120	-0,090	0,486	0,820
Produção de Leite-Litro	0,872	-0,163	0,923	-0,079	-0,115	0,978	0,150
Valor da Produção-R\$	0,899	-0,160	0,948	-0,048	-0,114	0,974	0,173
Nº Vacas Ordenhadas	0,947	-0,134	0,978	-0,093	-0,124	0,874	0,449
Pop Urbana	0,057	0,765	0,048	0,781	0,818	0,039	0,055
Pop Rural	0,737	-0,009	0,709	-0,167	-0,049	0,165	0,935

Com o que foi visto até aqui, é pertinente fazer uma classificação de cada fator, assim, pode-se chamar o Fator 1 dos anos 1991 e 2000 de Fator do Desenvolvimento Humano, já o Fator 2 dos mesmos anos pode ser dito de Fator Bovinocultor. Já no ano de 2010, o Fator 1 possui a mesma composição do Fator 2 dos outros anos e também pode ser dito de Fator do Desenvolvimento Humano, enquanto o Fator 2 pode ser interpretado como o Fator dos Derivados da Bovinocultura e o Fator 3 como o Fator do Rebanho Bovino. Para o ano 1991, o Fator 1 é responsável por explicar 44,33% da variância e o Fator 2 explica 33,56%. Em 2000 o primeiro fator responde por 47,58% da variância explicada e o Fator 2 por 34,31%. E em 2010, o Fator 1 explica 34,46% da variância e os fatores 2 e 3 são responsáveis por explicar 33,55% e 21,73 da variância, respectivamente.

Com isso, a análise fatorial foi substancial para poder afirmar que houve uma homogeneidade na separação dos fatores nos três anos estudados, de modo que, tanto as matrizes de correlação de Pearson quanto a análise fatorial, mostraram que não há correlação entre o IDHM e a pecuária leiteira no estado. As matrizes de correlação indicam ainda correlação negativa entre as variáveis da pecuária e do IDHM, que indica que à medida que uma cresce, a outra pode sofrer queda.

Seguindo com as análises, agora para verificar os municípios que mais se assemelham em relação à ligação entre o IDHM e a bovinocultura leiteira, procedeu-se com a análise de agrupamentos de ligação completa e distância euclidiana. Foi medida ainda a relação entre a matriz de similaridade e cofenética com o intuito de verificar o nível de ligação que há entre as duas e se o dendrograma apresentará uma boa reprodução dos grupos formados. O coeficiente de correlação cofenética, calculado para os três anos, mostrou-se satisfatório, uma vez que apresentou valores acima de 0,8 (Tabela 9).

Tabela 9. Coeficiente de correlação cofenética para os anos 1991, 2000 e 2010

<b>Ano</b>	<b>Coeficiente de Correlação Cofenética*</b>
1991	0,8894
2000	0,9025
2010	0,9683

\* Para ser satisfatório deve ser maior que 0,8.

Em 1991, a matriz de semelhanças e a matriz de correlação cofenética mostraram-se bem correlacionadas e geraram um coeficiente de correlação cofenética satisfatório de 0,8894, desse modo é pertinente seguir com a análise dos grupos formados pela distância euclidiana por ligação completa. Em relação ao número de grupos formados, três agrupamentos mostraram ser uma boa escolha quando analisados com a base de dados. Para este ano, a formação dos grupos seguiu de acordo com a produtividade leiteira, de modo que ficaram alocados no primeiro grupo 57 municípios, que foram caracterizados como os menores produtores de leite. No segundo grupo estão 14 municípios que fazem parte dos maiores produtores de leite e o terceiro grupo está composto pelos principais produtores de leite do estado. Vale observar que nos segundo e terceiro grupos estão 6 dos 7 municípios que compõem a região do alto sertão sergipano, esta região possui a particularidade de ser a principal bacia leiteira do estado.

O coeficiente de correlação cofenética para os dados de 2000 foi de 0,9025, afirmando também uma boa relação entre as matrizes de distância e a cofenética. Aqui, os municípios foram agrupados conforme produtividade leiteira e, também, oferece melhor agrupamento quando divididos em 3 grupos. No grupo 1 ficaram alocados 49 municípios os quais são menos relevantes em relação à produtividade leiteira. O segundo grupo relacionou 20 municípios com boa produtividade leiteira, mas os principais produtores ficaram no grupo 3, formando um cluster de 6 municípios. No último grupo, todos os componentes fazem parte do alto sertão sergipano.

No ano de 2010, foram formados 3 grupos, seguindo o mesmo critério dos outros anos, e cada grupo foi composto pelos municípios segundo sua importância na produção leiteira. A utilização da ligação completa para esse estudo também foi confirmada pelo coeficiente de correlação cofenética (0,9683). Com isso, os grupos ficaram da seguinte forma: o grupo 1 com 68 municípios, o grupo 2 com 4 municípios e o terceiro grupo com 3 municípios. Os dois últimos grupos ficaram compostos por cidades localizadas no alto sertão sergipano, com exceção de um único município que faz parte do médio sertão. Sergipe, assim como os demais estados nordestinos, segue a tendência de as bacias leiteiras ficarem localizadas principalmente nas regiões semi-áridas, sertão e agreste, movidos pela forte presença da agricultura familiar nessas áreas (FILHO, 2007).

Quando verificados os municípios que compõem cada grupo em cada ano, percebemos a forte tendência dos municípios do alto sertão sergipano em agruparem-se,



principalmente nos grupos 2 e 3. Essa região é conhecida por abranger as maiores bacias leiteiras do estado, e a análise de agrupamentos só vem a confirmar isso. Outra região que também se destacou no ramo da pecuária leiteira foi o médio sertão sergipano, presente entre os agrupamentos das maiores pecuárias leiteiras e, em seguida, o agreste central sergipano, que em 2000 agrupou todos os seus 6 municípios junto aos principais produtores de leite.

Para facilitar a visualização do que foi relatado em relação aos agrupamentos, as Figuras 2, 3 e 4 foram dispostas em dendrogramas apresentando os grupos que foram formados pelos municípios.

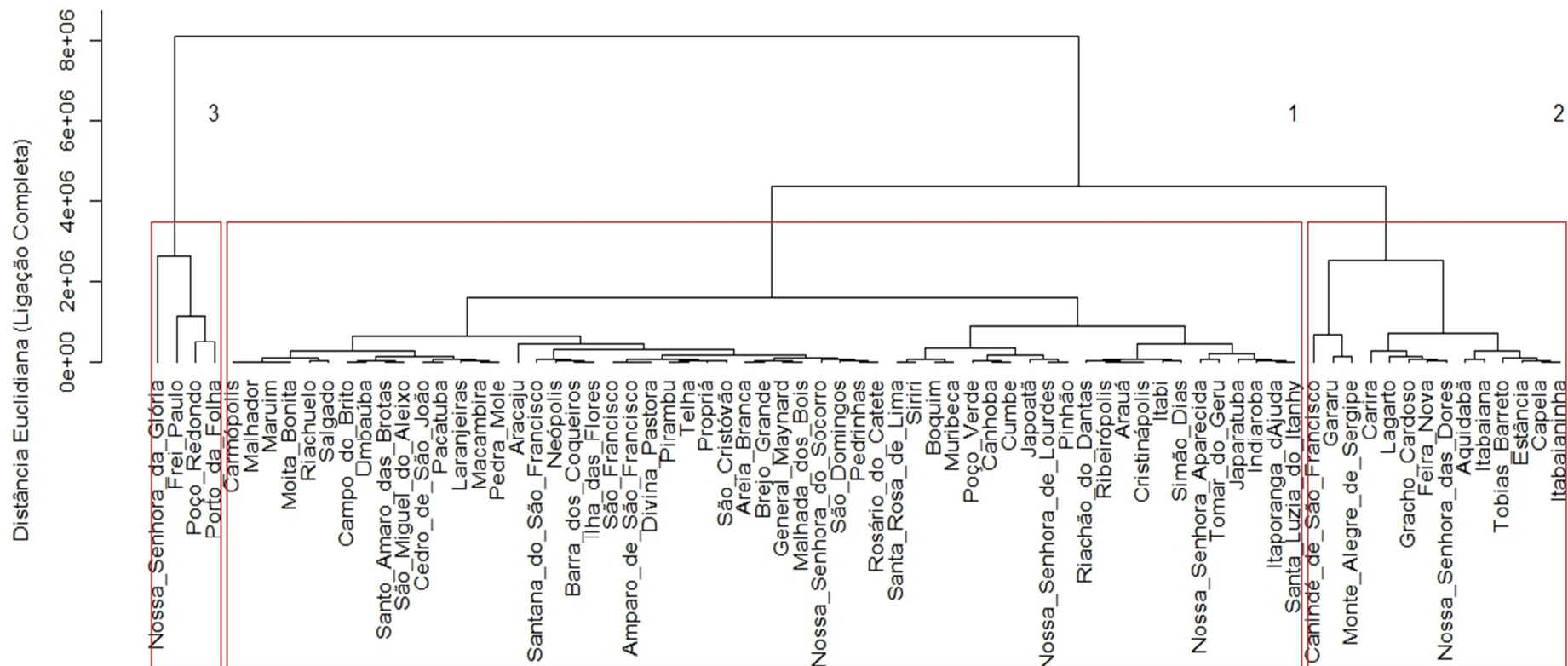


Figura 2: Dendrograma dos municípios de Sergipe em relação às variáveis dos IDHMs e da bovinocultura leiteira para o ano de 1991.

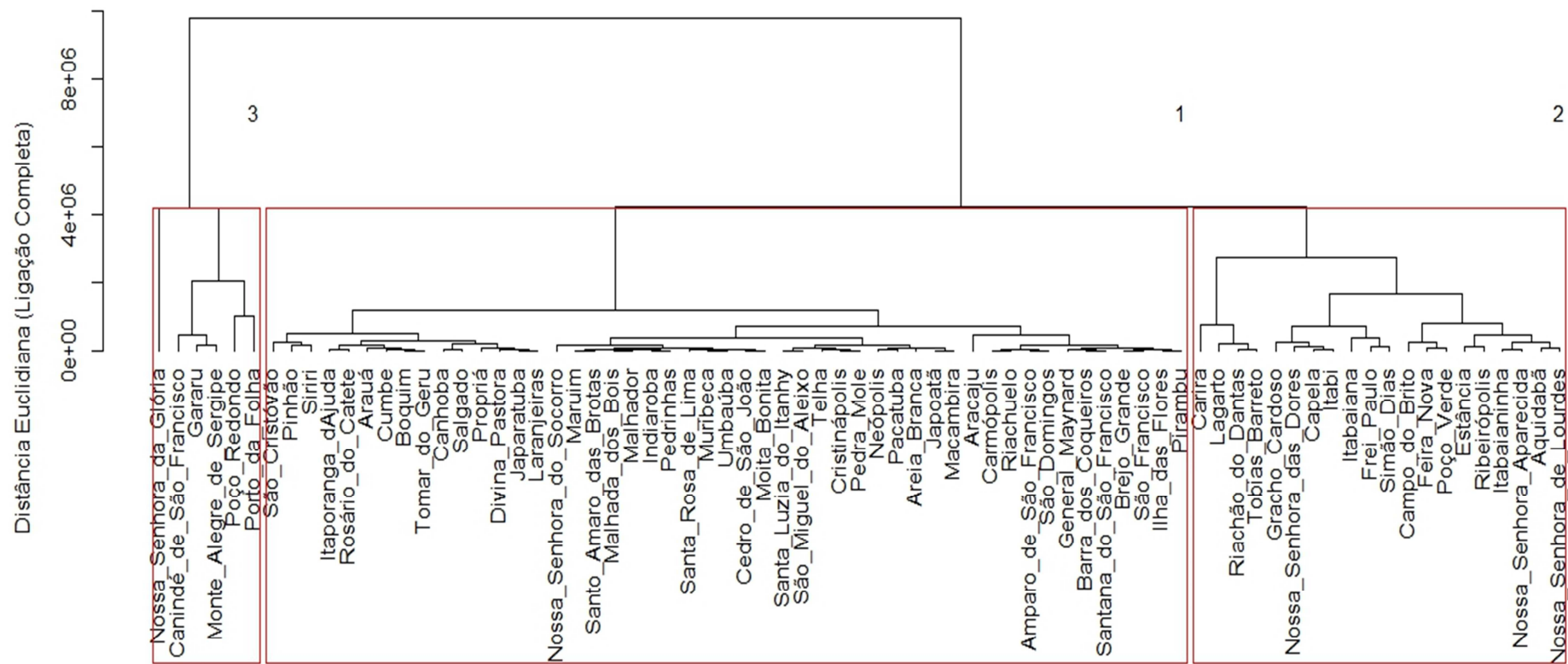


Figura 3: Dendrograma dos municípios de Sergipe em relação às variáveis dos IDHMs e da bovinocultura leiteira para o ano de 2000.

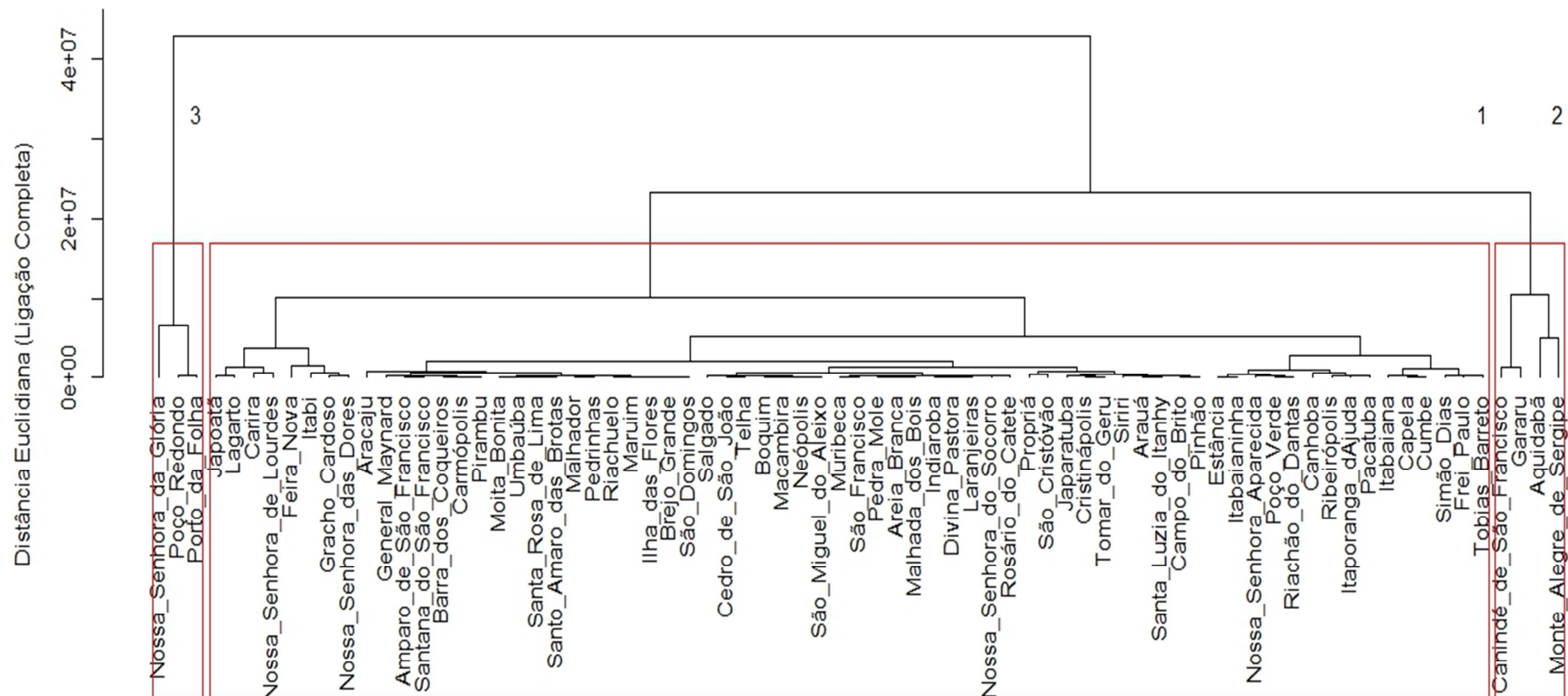


Figura 4: Dendrograma dos municípios de Sergipe em relação às variáveis dos IDHMs e da bovinocultura leiteira para o ano de 2010.

## 7 CONCLUSÕES

Fundamentado pela grande importância dos índices de desenvolvimento humano para a avaliação nos níveis de bem-estar social e pela importância da bovinocultura leiteira na economia local, a pesquisa ora desenvolvida teve por objetivo mostrar à sociedade se o crescimento da atividade leiteira nos municípios de Sergipe está refletindo no aumento dos IDHMs nas últimas três décadas. Para tanto, a partir dos dados provenientes das bases do IBGE e do PNUD, foram realizados métodos multivariados, por intermédio da análise fatorial e análise de agrupamentos.

Para a execução da análise fatorial foram necessárias uma série de testes que confirmassem sua viabilidade, e a primeira coisa que foi obtida foi a matriz de correlação de Pearson, nela foi revelada uma correlação fraca e negativa entre as variáveis do IDHM e da pecuária. O que significa dizer que essas variáveis não são apenas pouco correlacionadas, mas também pode-se dizer que à medida que uma cresce a outra pode sofrer decréscimo.

A separação das variáveis por fatores resultou em 2 fatores para os anos 1991 e 2000 e em 3 fatores para 2010. A composição desses fatores, nos três anos, era entre as variáveis do IDHM e a população urbana e entre as variáveis da pecuária e a população rural. Assim, como a análise fatorial viabiliza a identificação de variáveis correlacionadas dentro de um mesmo fator, confirmou-se aqui o que a matriz de correlação já havia anunciado, que o crescimento do desenvolvimento social local não está diretamente ligado ao crescimento da pecuária leiteira.

O passo seguinte foi a realização de agrupamentos que permitissem separar os municípios por características. Nesse sentido foram calculadas, nas três décadas estudadas, matrizes de distância euclidiana por ligação completa que se mostraram métodos apropriados através do cálculo do coeficiente de correlação cofenética, os quais assumiram valores maiores que 0,8. Em cada ano investigado foram separados 3 grupos, em cada grupo obteve-se os municípios com características similares entre si. De acordo com os agrupamentos, os municípios alocaram-se segundo sua importância na produção leiteira, de modo que estiveram no grupo 1 os menores produtores, no grupo 2 as localidades que possuem boa produção de leite e no terceiro grupo ficaram as principais bacias leiteiras do estado. Cabe ainda ressaltar, a importância que a pecuária leiteira possui para os municípios do alto sertão sergipano.

Por fim, percebeu-se ao longo da pesquisa, que o crescimento do desenvolvimento humano nos municípios sergipanos não está diretamente ligado à ascensão da pecuária leiteira, e que essa atividade econômica está fortemente ligada à população rural. Esta análise mostrou também que, não necessariamente um município que possui alto nível de produção leiteira tem o IDHM baixo, ou seja, a localidade pode estar entre os maiores produtores de leite e possuir IDHM alto, pois são fatores que independem da ocorrência do outro para que aconteça.

## BIBLIOGRAFIA

BARBETTA, P. A. **Estatística aplicada às ciências sociais**. 5. ed. rev. Florianópolis: UFSC, 2002.

BEZERRA, F. A. **Análise Fatorial**. In: CORRAR, L.; PAULO, E.; DIAS FILHO, J. **Análise multivariada para os cursos de administração, ciências contábeis e economia**. São Paulo: Atlas, 2007.

BUSSAB, W. de O.; MIAZAKI, E. S.; ANDRADE, D. F. **Introdução à análise de agrupamentos**. São Paulo: ABE, 1990.

CARVALHO, M. E. S.; SOUTO, P. H. **Da colonização ao século XXI: fragmentos geográficos da paisagem na bacia sergipana do Rio Vaza Barris**, 2011. Disponível em:

<[https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwi017Ca9s7JAhVFIpAKHa3VB8gQFggcMAA&url=http%3A%2F%2Fcascavel.ufsm.br%2Frevistas%2Fojs-2.2.2%2Findex.php%2Fgeografia%2Farticle%2Fdownload%2F7363%2F4402&usg=AFQjCNENL7g09C4s\\_QUZRUQ-aFR0C6vXnQ&bvm=bv.109332125,d.Y2I](https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwi017Ca9s7JAhVFIpAKHa3VB8gQFggcMAA&url=http%3A%2F%2Fcascavel.ufsm.br%2Frevistas%2Fojs-2.2.2%2Findex.php%2Fgeografia%2Farticle%2Fdownload%2F7363%2F4402&usg=AFQjCNENL7g09C4s_QUZRUQ-aFR0C6vXnQ&bvm=bv.109332125,d.Y2I)>. Acesso em: 02 fev. 2015.

CRUZ, I. C.; TOPA, M. A. **Análise multivariada como ferramenta de gerenciamento de fornecedores visando um relacionamento com vantagem competitiva**, 2009. Disponível em: <[http://www.coordest.ufpr.br/wp-content/uploads/2011/11/IvaneCruz\\_MarcosTopa.pdf](http://www.coordest.ufpr.br/wp-content/uploads/2011/11/IvaneCruz_MarcosTopa.pdf)>. Acesso em: 06 jan. 2015.

FIALHO, W. **A criação de gado de leite no Brasil**. *Revista Agropecuária*, 2012. Disponível em: <<http://www.revistaagropecuaria.com.br/2012/04/17/a-criacao-de-gado-de-leite-no-brasil/>>. Acesso em: 03 fev. 2015.

FILHO, O. M. C. et al. **Produção de leite em sistema agroecológico no semi-árido sergipano.** (2007). Disponível em: <<http://www.abagroecologia.org.br/revistas/index.php/rbagroecologia/article/view/6611/4916>>. Acesso em: 05 dez. 2015.

GUIMARÃES, J.R.S e JANNUZZI, P.M. **IDH, Indicadores sintéticos e suas aplicações em políticas públicas: uma análise crítica**, 2005. Disponível em: <<http://unuhospedagem.com.br/revista/rbeur/index.php/rbeur/article/view/136>>. Acesso em: 01 dez. 2014.

HAIR Jr., J.F.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E. & TATHAM, R.L. **Análise multivariada de dados.** 5.ed. Porto Alegre, Bookman, 2005.

IBGE. **Banco de Dados Agregados.** Sistema IBGE de Recuperação Automática - SIDRA. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 15 out. 2014.

IBGE. **Contas Regionais do Brasil.** Disponível em: <<http://www.ibge.gov.br/home/estatistica/economia/contasregionais/2013/default.shtm>>. Acesso em: 20 nov. 2015.

IBGE. **Pesquisa Pecuária Municipal – PPM.** Sistema IBGE de Recuperação Automática - SIDRA. Disponível em: <<http://www.sidra.ibge.gov.br/bda/pesquisas/ppm/default.asp>>. Acesso em: 20 nov. 2015.

KIELING, L. M. **O índice de desenvolvimento humano: adaptações metodológicas e práticas no Brasil**, 2014. Disponível em: <<https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/124693/Monografia%20da%20Luisa%20Michels%20Kieling.pdf?sequence=1&isAllowed=y>>. Acesso em: 15 jan. 2015.

MARQUES, A. F. **Aplicação da análise multivariada na infraestrutura e no desempenho das escolas públicas do Ensino Fundamental e Médio pertencentes ao**



**Núcleo Regional de Educação de Paranavaí**, 2010. Disponível em: <<http://periodicos.uem.br/ojs/index.php/ActaSciTechnol/article/view/1188/1188>>. Acesso em: 21 jan. 2015.

MENEZES, S. S. M. **O "saber-fazer apurado" do Requeijão do Sertão: a tradição e a cultura definham-se em Sergipe**. São Cristovão, 2012. Disponível em: <<http://www.cienciadoleite.com/?action=1&a=334&type=0>> Acesso em: 07 mar. 2015.

MEYER, A. S. **Comparação de coeficientes de similaridade usados em análises de agrupamento com dados de marcadores moleculares dominantes**, 2002. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/11/11134/tde-24072002-165250/pt-br.php>>. Acesso em: 05 dez. 2015.

OTTONI, C. **Indicadores sociais na formulação de políticas públicas federais brasileiras: teoria e prática**, 2006. Disponível em: <[http://repositorio.unb.br/bitstream/10482/4845/1/2006\\_Cristiano%20Ottoni.pdf](http://repositorio.unb.br/bitstream/10482/4845/1/2006_Cristiano%20Ottoni.pdf)>. Acesso em: 03 mar. 2015.

PNUD. **Atlas do Desenvolvimento Humano no Brasil**. Disponível em: <<http://www.pnud.org.br/>>. Acesso em: 15 out. 2014.

ROQUE, A. **Desenvolvimento como liberdade: uma aplicação dos conceitos de Amartya Sen à educação de adultos**, 2009. Disponível em: <[http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select\\_action=&co\\_obra=154986](http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=154986)>. Acesso em: 21 jan. 2015.

SEBRAE. **Produção de leite no Nordeste cresce no 2º trimestre**, 2014. Disponível em: <<http://www.infonet.com.br/economia/ler.asp?id=165057>>. Acesso em: 02 fev. 2015.

VARGAS, M. A. M. **Territórios de identidade nos territórios de planejamento**, 2011. Disponível em: <<http://anpege.org.br/revista/ojs-2.2.2/index.php/anpege08/article/viewFile/150/RAET5>>. Acesso em: 09 mar. 2015.

VICINI, L. **Análise Multivariada da Teoria à Prática**, 2005. Disponível em: <<http://w3.ufsm.br/adriano/livro/Caderno%20dedatico%20multivariada%20-%20LIVRO%20FINAL%201.pdf>>. Acesso em: 10 fev. 2014.

## **APÊNDICE A** – Algoritmo utilizado para o cálculo da Análise de Agrupamentos no R-project

```
# carrega a biblioteca e os dados
install.packages("vegan")
library(vegan)
idh.91=read.table(file.choose(),header=T)
attach(idh.91)

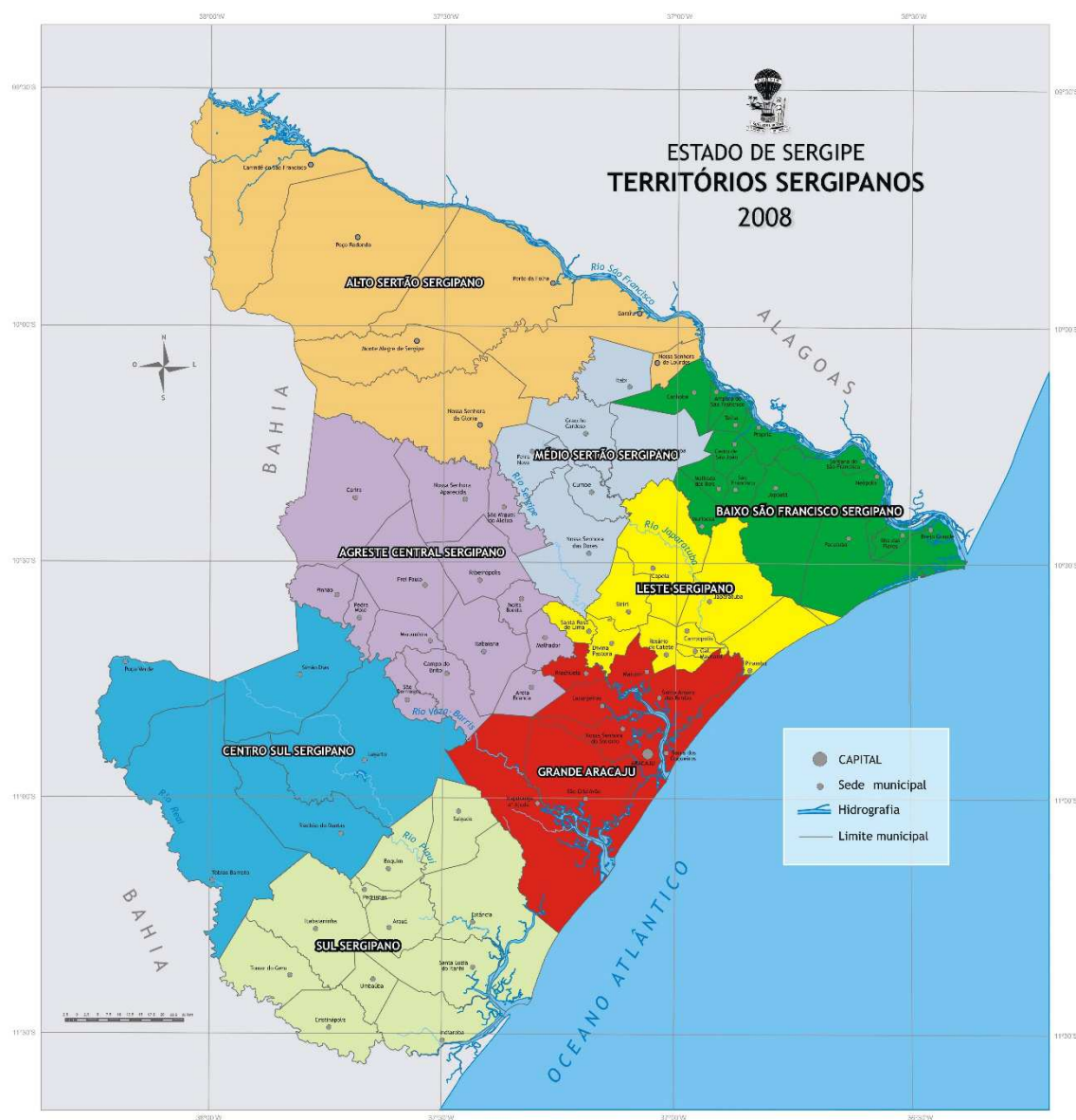
# calcula a matriz de distância, as ligações e plota o dendrograma
dis <- vegdist(idh.91,method="euclidean")
cluc <- hclust(dis, method="complete")
plot(cluc,hang=-1,main = "",xlab = "", ylab = "Distância Euclidiana (Ligação
Completa)")

# verifica que objetos foram classificados em cada grupo
grp<-cutree(cluc, 3)

# desenha no dendrograma os retângulos de cada grupo e os numera
r <- rect.hclust(cluc, 3)
text(cumsum(sapply(r,length)),
rep(mean(tail(unique(cluc$hei),2)), length(r)),
paste(unique(grp[cluc$ord]))))

# Cálculo do coeficiente de correlação cofenética
cor(dis,cophenetic(cluc))
```

## ANEXO A - Mapa dos territórios de planejamento do estado de Sergipe



Base Cartográfica: Atlas Digital Sobre Recursos Hídricos do Estado de Sergipe, 2004  
Elaboração: Gerência de Informações Geográficas e Cartográficas - GIGEC / SUPES / SEPLAN, 2008